Supplementary Note

RNAseq: statistical analysis.

We define the two reciprocal F_1 crosses between the inbred strains A and B as $A\times B$ and $B\times A$ where the maternal strain is given first. Suppose there are in total K_1 F_1 samples and K_2 inbred samples. We first count the number of RNA-seq reads that overlap exonic regions of each gene within each sample. For the ith ($i = 1, \dots, K_1$) F_1 sample, a gene has three counts: the total read counts (TReC) (m_i), the number of allele specific reads mapped to strain A (n_{iA}), and the number of allele-specific reads mapped to strain B (n_{iB}). For the lth ($l = K_1 + 1, \dots, K_1 + K_2$) inbred sample, a gene only has one count: the TReC (m_j). To remove lowly expressed genes, we exclude from our analysis any gene with a maximal TReC across all samples less than 50. We model the TReC by a negative binomial distribution and the allele-specific counts by a beta-binomial distribution to allw for possible over-dispersion commonly observed in RNAseq data at each gene.[1-3] For genes without enough allele-specific counts (i.e., average number of allele-specific counts in both $A\times B$ and $B\times A$ and both sexes are smaller than 5), we only modeled the TReC data. For the ith F_1 sample, we model n_{iB} with the following beta-binomial distribution:

$$p(n_{iB};n_i,\pi_i,\phi)=\binom{n_i}{n_{iB}}\frac{\prod_{k=0}^{n_{iB}-1}(\pi_i+k\phi)\prod_{k=0}^{n_i-n_{iB}-1}(1-\pi_i+k\phi)}{\prod_{k=1}^{n_i-1}(1+k\phi)}, \quad (1)$$

where $n_i = n_{iA} + n_{iB}$ is the total number of allele-specific reads. The parameter ϕ is an over-dispersion parameter. When $\phi = 0$, there is no over-dispersion and the beta-binomial distribution is simplified to a binomial distribution. We model the relationship between π_i, the expected proportion of allele-specific reads from strain B with paternal/maternal status as:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right)=(b_{0f}+b_{1f}x_i)I(female)+(b_{0m}+b_{1m}x_i)I(male) \quad (2)$$

where $x_i = 1$ if the sample is a $A\times B$ cross, otherwise$x_i = -1$. *I(event)* is an indicator function which equals 1 if the event is true, otherwise 0.

For TReC, we model m_l (where$l = K_1 + 1, \dots, K_1 + K_2$) via the negative binomial distribution with mean μ_l and overdispersion φ:

$$p(m_l;\mu_l,\varphi)=\frac{\prod_{k=1}^{m_l-1}(1+\varphi k)}{m_l!\varphi^{m_l}}\frac{(\varphi\mu_l)^{m_l}}{(1+\varphi\mu_l)^{m_l+\varphi^{-1}}}, \quad (3)$$

where $\log(\mu_l)=\beta_0+\beta_1\kappa_l+\beta_2 I(male)+\beta_3 dom+\beta_4 I(male)*dom+\eta_l$ with

$$dom = \begin{cases} 0 \text{ if sample } l \text{ is an } A \text{ or } B \text{ cross} \\ 1 \text{ if sample } l \text{ is an } A \times B \text{ or } B \times A \text{ cross} \end{cases}, \text{ and}$$

$$\eta_l = \begin{cases} 0 & \text{if sample } l \text{ is an } A \text{ cross} \\ b'_{0f} & \text{if sample } l \text{ is a } B \text{ cross} \\ -b_{1f} + \log\left\{1+\exp(b'_{0f}+b_{1f})\right\} - \log\{1+\exp(-b_{1f})\} & \text{if sample } l \text{ is an } A \times B \text{ cross} \\ \log\left\{1+\exp(b'_{0f}-b_{1f})\right\} - \log\{1+\exp(-b_{1f})\} & \text{if sample } l \text{ is a } B \times A \text{ cross.} \end{cases} \quad (4)$$

for female samples and

$$\eta_l = \begin{cases} \log\{1+\exp(-b_{1m})\} - \log\{1+\exp(-b_{1f})\} & \text{if sample } l \text{ is an } A \text{ cross} \\ b'_{0m} + \log\{1+\exp(-b_{1m})\} - \log\{1+\exp(-b_{1f})\} & \text{if sample } l \text{ is a } B \text{ cross} \\ -b_{1m} + \log\left\{1+\exp(b'_{0m}+b_{1m})\right\} - \log\{1+\exp(-b_{1f})\} & \text{if sample } l \text{ is an } A \times B \text{ cross} \\ \log\left\{1+\exp(b'_{0m}-b_{1m})\right\} - \log\{1+\exp(-b_{1f})\} & \text{if sample } l \text{ is a } B \times A \text{ cross} \end{cases}$$

for male samples.

In this model we account for κ_l, the library size (the total number of reads of sample l) which is important when modeling the total number of reads mapped to a given gene. We also consider sex effects which include both strain-specific and parent-of-origin-specific sex effects. In addition, sex specific dominant effects are also modeled. If the overdispersion parameter $\varphi = 0$, the negative binomial distribution reduces to a Poisson distribution.

The joint likelihhood of the combined F_1 and inbred samples is therefore:

$$L(b_{0f}, b_{0m}, b'_{0f}, b'_{0m}, b_{1f}, b_{1m}, \beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \varphi, \phi) = \prod_{i=1}^{K1} p(n_{iB}; n_i, \pi_i, \varphi) \prod_{l=1}^{K1+K2} p(m_l; \mu_l, \phi).$$

which we maximize for obtaining the maximal likelihood parameter estimates.

We test strain and parent-of-origin effects using likelihood ratio tests as follows:

$$\text{Strain effect}: \; H_0 : b_{0f} = b_{0m} = b'_{0f} = b'_{0m} = 0 \quad vs. \quad H_1 : b_{0f} \neq 0,\ b_{0m} \neq 0, b'_{0f} \neq 0 \text{ or } b'_{0m} \neq 0. \quad (5)$$

$$\text{Parent of origin effect}: \quad H_0 : b_{1f} = b_{1m} = 0 \quad vs. \quad H_1 : b_{1f} \neq 0 \text{ or } b_{1m} \neq 0. \quad (6)$$

For sex effects, we consider the following four tests:

$$\text{joint sex effects}: \quad H_0 : b_{0f} - b_{0m} = b'_{0f} - b'_{0m} = b_{1f} - b_{1m} = \beta_2 = \beta_4 = 0.$$

$$\text{(additive) strain specific sex effect}: \quad H_0 : b_{0f} - b_{0m} = b'_{0f} - b'_{0m} = 0.$$

$$\text{parent of origin specific sex effect}: \quad H_0 : b_{1f} = b_{1m}.$$

$$\text{dominant specific sex effect}: \quad H_0 : \beta_4 = 0.$$

Note that when testing for additive strain effects, we set $b_{0,f} = b_{0,m} = 0$ which is equivalent to $\pi_A = \pi_B = 0.5$ where π_A and π_B are the proportions of reads mapped to strain A and B. For autosomal genes, this assumption is reasonable. However, for chrX, this assumption may not be correct. One X chromosome is silenced in each female cell as a result of X chromosome inactivation and, in a F_1 mouse, the choice of which X chromosome is silenced may not be random and can be biased by alleles at the X-linked X controlling element (*Xce*).[4-5] Ignoring *Xce* effects could lead to the incorrect identification of many genes with apparent strain effects that are simply due to the *Xce* effect. To account for the *Xce* effect, we modified the method: for each F_1 female sample, we calculate τ_{iA} and τ_{iB} which are the averages of the proportions of reads that mapped to strain A and B across all genes on the X chromosome (except *Xist*, since it is expressed from the inactive X) and replaced b_0 in Equation (2) with $\log\left(\frac{\tau_{iB}}{\tau_{iA}}\right) + b_0$, and reset η_l in Equation (4) for female mice to:

$$\eta_l = \begin{cases} 0 & \text{if sample } l \text{ is an } A \text{ cross} \\ b'_{0,F} & \text{if sample } l \text{ is a } B \text{ cross} \\ \log\left\{1+\exp(\log(\frac{\tau_{iB}}{\tau_{iA}}) + b'_{0,F} + b_{1,F})\right\} & \text{if sample } l \text{ is an } A\times B \text{ cross} \\ -\log\{1+\exp(b_{1,F})\} + \log(2\tau_{iA}) & \\ \log\left\{1+\exp(\log(\frac{\tau_{iB}}{\tau_{iA}}) + b'_{0,F} - b_{1,F})\right\} & \text{if sample } l \text{ is a } B\times A \text{ cross.} \\ -\log\{1+\exp(-b_{1,F})\} + \log(2\tau_{iA}) & \end{cases}$$

and for male mice to:

$$\eta_l = \begin{cases} -\log\left\{1+\exp(-b_{1,F})\right\} + \log\left\{2\right\} & \text{if sample } l \text{ is an } A\times. \text{ cross} \\ b'_{0,M} - \log\left\{1+\exp(-b_{1,F})\right\} & \text{if sample } l \text{ is an B}\times. \text{ cross} \end{cases}.$$

We can then follow the same estimation and testing procedures for the strain and parent-of-origin effects. When modeling total read counts for genes with no allele specific counts, we find that when there exists no strain effect, the parent of origin effect is not identifiable. We modify the above models slightly to avoid this model identifiability issue. Details can be found in Zou et al. (submitted). Though compared to other studies, our RNA-seq data has a larger number of samples and higher sequencing depth, our test statistics are inflated and depart from asymptotic chi-square distributions. Therefore, we employ the genomic control[6] approach common to association mapping and adjust our test statistics by appropriate inflation factors. Our extensive simulations showed superior performance of the genomic control approach. To account for multipe testing across genes, we used the R package 'qvalue'[7] to estimate the q-value for each gene and for each test. For the strain effect test, because the majority of genes have strain effect, the q-values could be even larger than p-values. Therefore we

use a significance cutoff whereby both the p-value and q-value must be smaller than 0.05. The significance assement of imprinting effects are described in the next paragraph.

For each gene, we calculated its imprinting p-value by combining the imprinting p-values of three crosses using Fisher's method. We then calculated q-values based on the (meta-) imprinting p-values. One final modification was made to facilitate the identification of parent of origin effects. Since imprinting status appears to have spatial dependence such that two genes adjacent to each other tend to have the same imprinting status, we decided to model the (meta-) imprinting p-values by a Hidden Markov Model (HMM) to detect some weaker parent-of-origin effects by borrowing information from nearby genes. Specifically, we construct a HMM with two states: imprinted or non-imprinted. For the non-imprinted state, we assume the p-values follow a uniform distribution, and for the imprinting state, we assume the p-values follow a beta distribution beta(0.03,200). Denote the imprinted/non-imprinted states by I and N respectively. The transition probability of this HMM is set as $Pr(N \rightarrow N) = 0.99$, $Pr(N \rightarrow I) = 0.01$, $Pr(I \rightarrow N) = 0.10$, and $Pr(I \rightarrow I) = 0.90$. The input data of this HMM are the imprinting p-values for 11828 genes, and their genomic position. We apply the Vertibi algorithm to find the most likely underlying states chromosome by chromosome. Each of the final list of 98 imprinting genes has q-value smaller than 0.01 and/or is identified as imprinted by the HMM. Only 8 imprinting genes are identified by HMM but do not pass q-value cutoff. The imprinting p-values of these 8 genes range from 0.00024 to 0.0084, and thus they do have marginally significant imprinting effect, but not significant enough to survive multiple testing correction.

1. Sun W. A statistical framework for eQTL mapping using RNA-seq data. *Biometrics* 2012; **68**(1)**:** 1-11.

2. Langmead B, Hansen KD, Leek JT. Cloud-scale RNA-sequencing differential expression analysis with Myrna. *Genome Biol* 2010; **11**(8)**:** R83.

3. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; **26**(1)**:** 139-140.

4. Lyon MF. Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature* 1961; **190:** 372-373.

5. Cattanach BM. Control of chromosome inactivation. *Annu Rev Genet* 1975; **9:** 1-18.

6. Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999; **55**(4)**:** 997-1004.

7. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 2003; **100**(16)**:** 9440-9445.